

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2003, 08:44:13 ; Search time 42 Seconds
(without alignments)
2468.309 Million cell updates/sec

Title: US-10-010-227-3

Perfect score: 4055
Sequence: 1 MPAESTPQTLTKVLQAHV.....KAVPPTTNGEKKPLEW 778

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

1: A_GeneSeq_101002.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1666.5	41.1	644	22	AAU34234
2	1507.5	37.2	474	22	AAU34631
3	1479.5	36.5	466	22	AAU34428
4	1471	36.3	466	22	AAU35571
5	1450.5	35.8	466	22	AAU38424
6	1347.5	33.2	481	22	AAU31198
7	1330	32.8	553	22	AAU37666
8	1316	32.5	534	22	AAU37677
9	1300	32.1	461	23	ABP39047
0	1282	31.6	456	22	AAU36563

11	1282	31.6	456	22	AAU37335	Staphylococcus aur
12	1279.5	31.6	462	23	ABU48172	Listeria monocytrog
13	1138.5	28.1	460	15	ABU54216	L.lactis branched
14	1132.5	27.9	460	23	ABU54552	Lactococcus lactis
15	749	18.5	264	19	AAU77717	3-Isopropylmalate
16	688.5	17.0	245	22	AAU81974	S. epidermidis ope
17	637.5	15.7	423	22	AAU66358	Putative 3-isoprop
18	558.5	13.8	461	21	AAU29924	Arabidopsis thalia
19	558.5	13.8	469	21	AAU40226	Arabidopsis thalia
20	558.5	13.8	509	21	AAU29923	Arabidopsis thalia
21	558.5	13.8	509	23	ABU92924	Herbically activ
22	558.5	13.8	517	21	AAU40225	Arabidopsis thalia
23	550.5	13.6	443	21	AAU29925	Arabidopsis thalia
24	550.5	13.6	451	21	AAU40227	Arabidopsis thalia
25	542	13.4	428	21	AAU56307	Pseudomonas aerugi
26	528.5	13.0	418	21	AAU28433	Corn leuc subunit
27	523.5	12.9	418	21	AAU5394	Thermus thermophil
28	523.5	12.9	443	21	AAU28434	Corn leuc subunit
29	506	12.5	201	22	AAU38423	Salmonella typhi c
30	495	12.2	200	22	AAU35572	Haemophilus influe
31	493	12.2	173	19	AAU77716	3-Isopropylmalate
32	491	12.1	201	22	AAU34427	E. coli cellular p
33	482.5	11.9	191	15	AAU54217	L.lactis branched
34	482.5	11.9	191	23	ABU54553	Lactococcus lactis
35	482.5	11.9	780	22	AAU84333	Amino acid sequenc
36	481.5	11.7	780	20	AAU98947	Mitochondrial acou
37	472.5	11.7	787	22	ABU58512	Drosophila melanog
38	464.5	11.5	146	19	AAU79390	Staphylococcus aur
39	461.5	11.4	200	22	AAU36055	Klebsiella pneumon
40	442	10.9	683	22	ABU58511	Drosophila melanog
41	424	10.5	769	22	ABU60903	Drosophila melanog
42	422	10.4	190	22	AAU36564	Staphylococcus aur
43	403	9.9	732	22	ABU5728	Novel human diagn
44	396	9.8	789	23	ABU39041	Staphylococcus epi
45	393	9.7	892	22	ABU62277	Drosophila melanog
46	389.5	9.6	197	22	AAU31199	C glutaminc prote
47	389.5	9.6	197	22	AAU79772	Corynebacterium gl
48	389	9.6	226	21	AAU41138	Human ORFX ORF902
49	380.5	9.4	902	22	ABU61080	Drosophila melanog
50	375.5	9.3	193	23	ABU48173	Listeria monocytrog
51	375.5	9.3	304	21	AAU27662	Arabidopsis thalia
52	375.5	9.3	291	22	AAU63182	Corynebacterium th
53	370.5	9.1	211	21	AAU27663	Arabidopsis thalia
54	367.5	9.1	943	22	AAU91445	C glutaminc prote
55	366.5	9.0	805	22	AAU62827	Propionibacterium
56	365.5	9.0	868	21	AAU93301	Amino acid sequenc
57	364.5	9.0	901	22	AAU61624	S. epidermidis ope
58	364.5	9.0	914	22	ABU40363	Staphylococcus epi
59	363	9.0	919	16	AAU84337	Arabidopsis thalia
60	363	9.0	927	16	AAU85598	Arabidopsis thalia
61	359.5	8.9	869	21	AAU93285	Amino acid sequenc
62	357	8.8	891	22	AAU29353	Novel mar regulato
63	356.5	8.8	380	22	AAU96131	Putative 3-isoprop
64	354.5	8.7	889	23	ABU57134	Mouse ischaemic co
65	353.5	8.7	764	16	AAU84345	Melion aconitase en
66	352.5	8.7	891	16	AAU28261	Maize aconitase pa
67	352.5	8.7	891	16	AAU84338	Novel human diagno
68	344.5	8.4	160	22	AAU61975	S. epidermidis ope
69	340	8.3	110	23	ABU606700	Human ORFX protein
70	336	8.1	144	22	AAU53144	Propionibacterium
71	330	8.1	144	22	AAU53144	Rat wild-type IRP
72	329	8.1	963	23	AAU19852	Human wild-type IR
73	328.5	8.1	963	22	AAU19851	Human wild-type IR
74	326.5	8.0	117	22	AAU53143	Propionibacterium
75	326	8.0	848	23	ABU53982	Lactococcus lactis
76	325	8.0	900	23	ABU48185	Listeria monocytrog
77	324	8.0	106	23	ABU32040	Human isomerase-11
78	318.5	7.9	139	23	ABU32235	Human isomerase-11
79	312.5	7.7	639	21	AAU74439	Neisseria meningit
80	309.5	7.6	639	21	AAU74440	Neisseria meningit
81	307	7.6	853	22	AAU35942	Helicobacter pylori
82	303	7.5	853	22	AAU35770	Helicobacter pylori

84 301 7.4 639 21 AAY74438
85 292 7.2 86 23 ABP08386
86 284.5 7.0 214 21 AAG27664
87 271 6.7 814 22 AAG19229
88 264.5 6.5 120 22 AAG81461
89 250 6.2 127 21 AAY28435
90 239 5.9 119 19 AAY38591
91 238 5.9 70 23 ABP04569
92 212.5 5.2 268 18 AAW20584
93 209 5.2 78 19 AAW79391
94 209 5.2 263 21 AAY28441
95 208.5 5.1 195 21 AAY28436
96 207.5 5.1 533 23 ABB89994
97 207 5.1 205 22 ABG05731
98 204.5 5.0 553 22 AAB79520
99 204.5 5.0 557 22 AAB79519
100 198.5 4.9 203 18 AAW20430

ALIGNMENTS

RESULT 1
AAU34234
ID AAU34234 standard; Protein; 644 AA.
XX
AC AAU34234;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #510.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
FN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206948P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-25931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52093.
XX
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5730; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 644 AA;

Query Match 41.1%; Score 1666.5; DB 22; Length 644;
Best Local Similarity 48.7%; Pred. No. 2.3e-139;
Matches 352; Conservative 92; Mismatches 198; Indels 81; Gaps 10;

QY 9 QTLYDKVLQAHVDEKLDGTVLLYIDRHLVHEVTSPOAFEGLRNAGRKVRPDCDCTLATTD 68
DB 3 QTLFDKVMNRHVLGKLEPQLLYIDLHLIHEVTSPOAFEGLRNQRKLRPDLTFATLD 62
QY 69 HNVPTTSRKALKDIASEFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRGQIVHVIGPEQG 128
DB 63 HNVPTI-----DIFN-IKDEIANKQITTLQNAIDFGVHIFDMGSDSQGVHVMVGPETG 115
QY 129 FTLPPTVCGDSHTSTHGAFALAFGIGTSEVEHVLATQCLITKRSKNMIRIQVDGELAP 188
DB 116 LTQPGKTI VCGDSHTATGAFALAFGIGTSEVEHVFATQTLWOTKPKNLKIDINGTLPT 175
QY 189 GVSKDVLVHAIGTIGTAGTGAVIEFGSVIRSLSEARMSICNMSIEGGARAGWAPD 248
DB 176 GVAKDILHLIKTYGVDFGTGVALEFTGETIKNLSMDGRMTICNMAIEGAKYGIIOPD 235
QY 249 EITFEYLRGRPLAPKYDSPWHKATQVMKNLQSDPGAKYDIDVDIADKDI VPTLTWGTSP 308
DB 236 DITFEYVKGPFADNF-----AKSVKWRRELYSDDDAIDFDRVIELDVSTLSPQVTWGTNP 290
QY 309 EDVVPITGVVPDPETFAATEAKKADGRMLQYMGLKAGTFMEDIPVDKVFISGCTNSRIED 368
DB 291 EMGVNFSEPPF-----EISDINDQRAYDMGLEPGQKAEDIDLGYVFLGSGCTNARLSD 343
QY 369 LRAAA VVKGRKAPNVKSAWVPGSGLVKTAOAEELGDKIFEEAGFEWREAGCSMCLGM 428
DB 344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKREAEKGLDITTFKNAGFEWREPGCSMCLGM 402
QY 429 NPDI LAPQERCASTSNRNFEGRCAGGRTHLMSVMAAAAGIVGKLADVRKLTIDYKASPH 488
DB 403 NPQVPEGVHCASSTNENFEGRCQGKARTHLVSPMAAAAIAHGFVDVRKV-----454
QY 489 IAAQKSTVTVKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSGAGLPKFTILKG 548
DB 455 -----VXMAAIKPTTYKG 469
QY 549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDFFVLNKEYRKAS 608
DB 470 KIVPLFNDNIDTDQIIPKVLKRIKSGSGFGPFAFDEWRYLPDGSNDPDPFNKPKYKGS 529
QY 609 ILVCTGANFGCGSREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGMLPIPIKDAQIE 668
DB 530 ILI-TGDNFCGSSREHAALWKDYGPHEIIAGSFSDFIYNNCTKNAMLPVLEKNAR-E 587
QY 669 AIAAARAGKEIEVDLPNQLIKNATGETICTIFEVEEPRKHLVNGLDDIGITMQMEDKIA 728
DB 588 HLAKYV-----EIEVDLPNQTV--SSPKSFHFEIDETWKNKLVNGLDDIAITLQYESLIE 641
QY 729 EFE 731
DB 642 KYE 644

RESULT 2
AAU33631
ID AAU33631 standard; Protein; 474 AA.

```
XX AC AAU33631;
XX DT 14-FEB-2002 (first entry)
XX DE Pseudomonas aeruginosa cellular proliferation protein #75.
XX KM Antisense; prokaryotic cellular proliferation protein;
XX KM antibiotic; antibacterial; drug design.
XX OS Pseudomonas aeruginosa.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 22-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS51490.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 5127; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 474 AA;
XX
XX Query Match 37.2%; Score 1507.5; DB 22; Length 474;
XX Best Local Similarity 62.5%; Pred No. 2,1e-125;
XX Matches 295; Conservative 61; Mismatches 113; Indels 3; Gaps 2;
```

```
QY 129 FTLPCTTVCCSDSTHTGAFGALAFGIGTSEVEHVLAATQCLITKRSNNRIQVDELAP 188
DB 123 ATLPGMTVWCSDSHSTHGAFGALAHGIGTSEVEHVLAATQCLVAKKMKMLKVXEGRLPA 182
QY 189 GVSSKDVVLAHAGIIGTAGGAGVIEFGSVIRSLSMERAKSICNMSIEGAGAMVAPD 248
DB 183 GVTAKDVLAVIGRIGTAGGAGHAIEFRGSAIRDSIEGRNITCNMSIEAGARGLVAD 242
QY 249 EITFEYLKGRPLAKPYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDANDIVPTLTWGTSP 308
DB 243 OKTIDYVKGRRFAP--SABQWDQAACWQGLVSDADARFDTVVELDAAQIKRQVSWGTSP 300
QY 309 EDVVPITGVVPDPETPATEAKKADGRMLQYMGIKAGTPMEDIPVDKFISSCTNSRIED 368
DB 301 EMVLAVDQNVDPAPRESPPIRGSIERALKYMGILRPNOAIDIDLVFISSCTNSRIED 360
QY 369 LRAAAVAVKGRKKAPNVASAWVVPQSGLVKTQAEBEGLDKTFEEAGFPMRAGSCMCLM 428
DB 361 LRAAAEVARGKRVAAATTKQALVPPQSGLVKQAEKEGLDRIFIEAGFPMRBPQSCMCLAM 420
QY 429 NPDILAPQERCASTSNRNFEGRQAGGRTHLMSPVMAAAGIVGKLADVRL 480
DB 421 NPDRLESGEHCASTSNRNFEGRQAGGRTHLVSPMAAAAANGRFIDVREL 472
XX
XX RESULT 3
XX ID AAU34428
XX AC AAU34428 standard; Protein; 466 AA.
XX AC AAU34428;
XX DT 14-FEB-2002 (first entry)
XX DT 27-SEP-2001.
XX DE E. coli cellular proliferation protein #9.
XX DE Antisense; prokaryotic cellular proliferation protein;
XX KM antibiotic; antibacterial; drug design.
XX OS Escherichia coli.
XX PN WO200170955-A2.
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52287.
XX PI Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52287.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX PS Example 3; Seq ID No 10021; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
```



```

Db      355  LRAAAAWMKGKKADNNKRIILVPGSSGLVKEQAEYEGJDKFLFIAGAEWRNPGSCMIGM 414
Qy      429  NPDLIAQOERCASSTNRNFEGRQAGRTHMSPYMAAAGIVGLADVRKLT 481
Db      415  NDRILGWERCASSTNRNFEGRQGRNGRTHLVSPMAAAGVFGKFDIVDT 467

RESULT 5
AAU38424
ID      AAU38424 standard; Protein; 466 AA.
XX
XX      AAU38424;
XX
XX      14-FEB-2002 (first entry)
XX
XX      Salmonella typhi cellular proliferation protein #315.
DE
XX      Antisense; prokaryotic cellular proliferation protein;
XX      antibiotic; antibacterial; drug design.
XX
XX      Salmonella typhi.
XX
XX      WO200170955-A2.
XX
XX      27-SEP-2001.
PD
XX      21-MAR-2001; 2001WO-US09180.
XX
XX      21-MAR-2000; 2000US-191078P.
PR
XX      23-MAY-2000; 2000US-206848P.
PR
XX      26-MAY-2000; 2000US-207727P.
PR
XX      23-OCT-2000; 2000US-242578P.
PR
XX      27-NOV-2000; 2000US-253625P.
PR
XX      22-DEC-2000; 2000US-257931P.
PR
XX      16-FEB-2001; 2001US-269308P.
XX
XX      (ELIT-) ELITRA PHARM INC.
XX
XX      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX      Yamamoto RT, Xu HH;
XX      WPI; 2001-611495/70.
XX      N-PSDB; AAS56283.
XX
XX      New polynucleotides for the identification and development of
XX      antibiotics, comprise sequences of antisense nucleic acids -
XX
XX      Example 3; Seq ID No 14017; 51bp; English.
XX
XX      The invention relates to antisense inhibitors of genes essential to
XX      prokaryotic cellular proliferation, their use in identifying the
XX      genes, their use in the discovery of novel antibiotics, the essential
XX      genes, themselves and the encoded proteins. The prokaryotes used are
XX      Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX      pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX      invention is also useful for the identification of potential new targets
XX      for antibiotic development. The antisense nucleic acids can also be used
XX      to identify proteins used in proliferation, to express these proteins,
XX      and to obtain antibodies capable of binding to the expressed proteins.
XX      The proteins can be used to screen compounds in rational drug discovery
XX      programmes. The antisense nucleic acid sequence is also useful to screen
XX      for homologous nucleic acids which are required for cell proliferation in
XX      a wide variety of organisms. The present sequence represents an
XX      essential prokaryotic cellular proliferation protein.
XX      Note: The sequence data for this patent did not form part
XX      of the printed specification, but was obtained in electronic
XX      ftp.wipo.int/pub/published_pct_sequences.
XX
XX      Sequence 466 AA;
XX
XX      Query Match 35.8%; Score 1450.5; DB 22; Length 466;

```

Best Local Similarity 60.6%, Pred. No. 2.5e-120;
Matches 286; Conservative 63; Mismatches 114; Indels 9; Gaps 4;

```
OY      9 QTLVDRKVLQAHNVADKKLDGTVLLYLIDRLVHEVTSPQAPEGILRNAGRKVRRPDCLATTD 68  
       3 :|::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db     3 KLVYEKLFDPAHVVFAPARPETPLLVIDRLVHEVTPSQAFDGLRAHHRVVRPGKTFATMD 62  
  
OY      69 HNVPTTSRALKDLASFIKEDDSTQCCTLLENNVEKESVTVFGLSDKRQGIIVHYIGPEQG 128  
       6 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db     63 HNVSOTL-----KDINA--SGEMARIQMCELIKNCFEFEVELYDLNHPYXGIVHWVGPEQX 116  
  
OY     129 FFLPCTVWCGSHSTHGAFGALFGLGTSEVEHVLTATOLLTKRSNMRIQVDGELAP 188  
       ||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db    117 VTLPEMKVLCGSHTAHXGARFGALKFEGITSEVEHVLTATOLLKOGRAKTMKI EVTGNAA 176  
  
OY     189 GVSSKDVVLHAIGIIGTAGTGAVIEFCGSVIRSHSMEARMSICNMSIEGARAGMVA 248  
       |177 GTTADIDYLAIIKTGKSAGCTGHVVEFPFGDAIRALSNEGRTLICMLAEKAKAGLVAP 236  
  
OY     249 EITFEYLKGRPLAPKYDSPENHKATQYWKLQSDPGAKYDIDEIDA KDIPTLTWTGSTP 308  
       2237 ETTFPVYVKGRHLAPK--GRDFEAIVEYWKTLKTTDGATFDIVTLRABEIAPQVTWGTP 294  
  
OY     309 EBYVITGVVPDEPFATEAKADGRMLQYMGIKAGTPMEDIVVDKVFISCTNSRIED 368  
       295 GGVIIVSTDIIPDPASFSDPVERASA EKALAYMGLQPGLPTDVAIDKVFISCTNSRIED 354  
Db    369 LRAAAYVKGRRKAPNVNASAMVPPSGLVKTQAAEEBGLDKIFEEAGFPMWRAGSMCLGM 428  
       355 LRAAAEVAKKRRKVAPEGV-ALVPPSGFKQAABEBGLDKIFIEAGFEMWRLPSCMCLAM 413  
  
OY     429 NPDIAPOEBCASTSNRNFEGRQGAAGRTHIMSEPMAAAGIVGKLADVRL 480  
       414 NNDRLNPBERCASTSNRNFEGRQGRGRTLHVSPMAAAA AVTHGFDIRSI 465  
Db  
  
RESULT 6  
AAG91198  
ID AAG91198 standard; Protein; 481 AA.  
XX  
AC AAG91198;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum protein fragment SEQ ID NO: 4952.  
XX  
KM Corynebacterium; amino acid synthesis; vitamin; saccharide;  
KW organic acid synthesis.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN EPI108790-A2.  
PD 20-JUN-2001.  
XX  
PF 18-DEC-2000; 2000EP-0127688.  
XX  
PR 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280988.  
XX  
PA (KYOW ) KYOMA HAKKO KOGYO KK.  
XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX  
XX WPI; 2001-376931/40.  
DR N-PSDB; AAA66417.  
XX  
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
XX mutation point of a gene, measuring expression of a gene, analysing  
XX expression profile or pattern of a gene and identifying homologous gene
```

XX PS Claim 17; SEQ ID NO: 4952; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein

XX CC sequences from the Corynebacterium glutamicum. These

XX CC are useful for identifying the mutation point of a gene derived from a

XX CC mutant of corynebacterium, measuring expression amount and

XX CC analysing the expression profile or expression pattern of a gene derived

XX CC from Corynebacterium, and identifying a homologue of a gene derived

XX CC from corynebacterium. Corynebacterium bacteria are useful for producing

XX CC amino acids, nucleic acids, vitamins, saccharides and organic acids,

XX CC particularly L-lysine. The present sequence is a protein described

XX CC in the exemplification of the invention.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from the

XX CC European Patent Office.

XX SQ Sequence 481 AA;

Query Match 33.2%; Score 1347.5; DB 22; Length 481;

Best Local Similarity 57.6%; Pred No. 4.1e-111;

Matches 273; Conservative 61; Mismatches 127; Indels 13; Gaps 6;

Qy 10 TLYDKVLQAVVDEKLDGTV-LLYIDRLVHVETSPQAFEGRLNAGKVRPDPCTLATTD 69

Db 15 TLAEKVWRDHVVKSGENGEPDLYIDQLLHEVTSQAFDGLRMTGRKLHPHLLIATED 74

Qy 69 HNVPTTSKALKDIASFKEEDP--SRTQCVTLLENVKEFGVYFGLSDKROGIVHVIGPE 126

Db 75 HNVPTTEGINT----GSLLEINDKISRLQVSTLRDNCSEFGVRLHPMGDVRQGIHVTVGPQ 130

Qy 127 QGFTLPPTVCGDSHTSTHGAFALFGIGTSEVHVLTATQCLITKRSKNMIRIQVDGEL 186

Db 131 LGATPGMTIVCGDSHTSTHGAFGSAFGIGTSEVHVWATQTLPLKPKTMAIEVTGEL 190

Qy 187 APGVSSKDVLLHAIGTAGTGAVIEFCGVSIRLSWEARMSICNMSIEGGARAGMVA 246

Db 191 QFVGSSKDLILAIKIGTGGQGVYVLEVRGEAIRKMSMDARMTMCNMSIEAGARAGMIA 250

Qy 247 PDEITFEYLKGRPLAPKYPDSPEWHKATQYKQLQSPGAKYDIDVFIDAKDIVPTLTWGT 306

Db 251 PDQTTFYVEGREMAPK--GADWDEAVAYWKTLPTEGATFDKVKVLDGSAITPFTWGT 308

Qy 307 SPEDVVPIGVVPDPFTFATEAKKADGRRLQYMLGKAGTPEMDIPVDKVFIGSCNRSRI 366

Db 309 NFGQGLPIGESVSPEDFTDNDKAAERKALQYMDLVPGTFLRDIKIDTVFLGSCINARI 368

Qy 367 EDLRAAAAVVGRKKAPNVKSAVVPGSGLVKTQAEEGLDKIFEEAGFEWRAGCSMCL 426

Db 369 EDLQIAADILKGHIADGMR--MNVVPSSTWIKQEAELGJDKIFTDAGAEMRTAGCSMCL 427

Qy 427 GNPDPILAQERCASSTNRFEGROGAGGTHLMSPVMAAAGLVGKL---ADV 477

Db 428 GNPDLQKPGERSASTSNRNFEGRQGGGRTHLVSPAVAAATAIRGLTSSPADI 481

RESULT 7

AAB79766

ID AAB79766 standard; Protein; 553 AA.

XX AC AAB79766;

XX DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:266.

XX KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;

XX KW fine chemical production; microorganism; organic acid; nucleoside;

XX KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;

XX KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;

XX KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.

XX OS Corynebacterium glutamicum.

XX WO200100843-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB00923.

XX 25-JUN-1999; 99US-0141031.

XX 01-JUL-1999; 99DE-1030476.

XX 02-JUL-1999; 99US-0142101.

XX 08-JUL-1999; 99DE-1031415.

XX 08-JUL-1999; 99DE-1031418.

XX 08-JUL-1999; 99DE-1031419.

XX 08-JUL-1999; 99DE-1031420.

XX 08-JUL-1999; 99DE-1031424.

XX 08-JUL-1999; 99DE-1031428.

XX 08-JUL-1999; 99DE-1031434.

XX 08-JUL-1999; 99DE-1031435.

XX 08-JUL-1999; 99DE-1031443.

XX 08-JUL-1999; 99DE-1031453.

XX 08-JUL-1999; 99DE-1031457.

XX 08-JUL-1999; 99DE-1031465.

XX 08-JUL-1999; 99DE-1031478.

XX 08-JUL-1999; 99DE-1031510.

XX 08-JUL-1999; 99DE-1031541.

XX 08-JUL-1999; 99DE-1031573.

XX 08-JUL-1999; 99DE-1031592.

XX 08-JUL-1999; 99DE-1031632.

XX 08-JUL-1999; 99DE-1031634.

XX 08-JUL-1999; 99DE-1031636.

XX 09-JUL-1999; 99DE-1032125.

XX 09-JUL-1999; 99DE-1032126.

XX 09-JUL-1999; 99DE-1032130.

XX 09-JUL-1999; 99DE-1032186.

XX 09-JUL-1999; 99DE-1032206.

XX 09-JUL-1999; 99DE-1032227.

XX 09-JUL-1999; 99DE-1032228.

XX 09-JUL-1999; 99DE-1032229.

XX 09-JUL-1999; 99DE-1032230.

XX 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032926.

XX 14-JUL-1999; 99DE-1032928.

XX 14-JUL-1999; 99DE-1033004.

XX 14-JUL-1999; 99DE-1033005.

XX 14-JUL-1999; 99DE-1033006.

XX 12-AUG-1999; 99US-0148613.

XX 27-AUG-1999; 99DE-1040764.

XX 27-AUG-1999; 99DE-1040765.

XX 27-AUG-1999; 99DE-1040766.

XX 27-AUG-1999; 99DE-1040832.

XX 31-AUG-1999; 99DE-1041378.

XX 31-AUG-1999; 99DE-1041379.

XX 31-AUG-1999; 99DE-1041380.

XX 31-AUG-1999; 99DE-1041394.

XX 31-AUG-1999; 99DE-1041396.

XX 03-SEP-1999; 99DE-1042076.

XX 03-SEP-1999; 99DE-1042077.

XX 03-SEP-1999; 99DE-1042079.

XX 03-SEP-1999; 99DE-1042086.

XX 03-SEP-1999; 99DE-1042087.

XX 03-SEP-1999; 99DE-1042088.

XX 03-SEP-1999; 99DE-1042095.

XX 03-SEP-1999; 99DE-1042124.

XX 03-SEP-1999; 99DE-1042129.

XX 09-MAR-2000; 2000US-0187970.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-137957/14.

XX N-PSDB; AAF71885.

PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
PS Claim 20; Page 545-547; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 553 AA;
Query Match 32.8%; Score 1330; DB 22; Length 553;
Best Local Similarity 57.9%; Pred. No. 1.9e-109;
Matches 267; Conservative 60; Mismatches 124; Indels 10; Gaps 5;
QY 10 TLVYKQLQAAHVDEKLDGTV-LTYIDRLVHEVTSPOAFEGLRNAGRVRPDTLTATTD 68
DB 15 TLAEKVRBHVSVSGENGEPDLATIDQLLHEVTSPOAFDELKRMGRKLRLHELHATED 74
QY 69 HNVPTTSRKALKIASFIKEDD--SRQCVTLENNVEFGVTVYFGLSDKROGIYHVIGPE 126
DB 75 HNVPTGIGIKT---GSLLEINDKISRLOVSTLRNCEFEFVRLHPMGDVROGIYHTVGPQ 130
QY 127 QGFTLPGTTVCGSHSTHGAFGALFGIGTSEVHVALQOCLITRSKMMRLOVDEEL 186
DB 131 LGATQGMATVCGSHSTHGAFSGSMFPGIGTSEVHVMATQTLPLKPKTMALEVTSEL 190
QY 187 APGVSSKDVVLAHIGLIGTAGTGAVIEFCGSVIRSLSEMARMSICNMSIEGARAGMVA 246
DB 191 QRGVSSKDLILATIAKIGTGGGQYVLEFGEALRKMSMDARMCMNSITAGARAGMIA 250
QY 247 PDEITEYLKGRPLARYDSPEWHKATQYMKNTLOSPQAKYDIDVFIDAKDIVPTLTWGT 306
DB 251 PQTITDYVEGREGMAPK--GADWDEAVAYWKTLPTEDEGATDKVVEIDGSLTFPIWTGT 308
QY 307 SPEDVVPITGVNVDPEPFATEAKKADRRMLQVNGLAKGTMEDIPVDKVIIGSTNSRI 366
DB 309 NRGQGLPLGESVSPEDFTTNDNDAAAEKALQYMDLVPGPLRDIKIDTVELSGCTNARI 368
QY 367 EDLRAAAAVVGRKKAPNVKSAMVVPSSGLVKTQAESEGLDKIFEEAGFEWREAGCSWCL 426
DB 369 EDLQIAADILKGRKIADGMR--MNVVPESTWIKQEAELGLDKIFTDAGAEWRTAGCSWCL 427
QY 427 GNNPDLIAFOERCASISNNRNFEGROGAGGRTHLMSPVMAAA 467
DB 428 GNNPDQLKPGERSAFTSNRNFEGROGGRTHVSPAVAAA 468
RESULT 8
ID AAB79767 standard; Protein; 534 AA.
XX AAB79767;
XX AAB79767;
XX 30-APR-2001 (first entry)
XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:268.
XX
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
XX fine chemical production; microorganism; organic acid; nucleoside;
XX nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
XX lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
XX carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX Corynebacterium glutamicum.
XX
XX MOZ00100843-AA2.

XX 04-JAN-2001.
PD 23-JUN-2000; 2000KO-IB00923.
XX
XX 25-JUN-1999; 99US-0141031.
XX 01-JUL-1999; 99DE-1030476.
XX 02-JUL-1999; 99US-0142101.
XX 08-JUL-1999; 99DE-1031415.
XX 08-JUL-1999; 99DE-1031418.
XX 08-JUL-1999; 99DE-1031419.
XX 08-JUL-1999; 99DE-1031420.
XX 08-JUL-1999; 99DE-1031424.
XX 08-JUL-1999; 99DE-1031428.
XX 08-JUL-1999; 99DE-1031434.
XX 08-JUL-1999; 99DE-1031435.
XX 08-JUL-1999; 99DE-1031443.
XX 08-JUL-1999; 99DE-1031453.
XX 08-JUL-1999; 99DE-1031457.
XX 08-JUL-1999; 99DE-1031465.
XX 08-JUL-1999; 99DE-1031478.
XX 08-JUL-1999; 99DE-1031510.
XX 08-JUL-1999; 99DE-1031541.
XX 08-JUL-1999; 99DE-1031573.
XX 08-JUL-1999; 99DE-1031592.
XX 08-JUL-1999; 99DE-1031632.
XX 08-JUL-1999; 99DE-1031634.
XX 08-JUL-1999; 99DE-1031636.
XX 09-JUL-1999; 99DE-1032125.
XX 09-JUL-1999; 99DE-1032126.
XX 09-JUL-1999; 99DE-1032130.
XX 09-JUL-1999; 99DE-1032186.
XX 09-JUL-1999; 99DE-1032206.
XX 09-JUL-1999; 99DE-1032227.
XX 09-JUL-1999; 99DE-1032228.
XX 09-JUL-1999; 99DE-1032229.
XX 09-JUL-1999; 99DE-1032230.
XX 14-JUL-1999; 99DE-1032922.
XX 14-JUL-1999; 99DE-1032926.
XX 14-JUL-1999; 99DE-1032928.
XX 14-JUL-1999; 99DE-1033004.
XX 14-JUL-1999; 99DE-1033005.
XX 14-JUL-1999; 99DE-1033006.
XX 12-AUG-1999; 99US-0148613.
XX 27-AUG-1999; 99DE-1040764.
XX 27-AUG-1999; 99DE-1040765.
XX 27-AUG-1999; 99DE-1040766.
XX 31-AUG-1999; 99DE-1040832.
XX 31-AUG-1999; 99DE-1041378.
XX 31-AUG-1999; 99DE-1041379.
XX 31-AUG-1999; 99DE-1041380.
XX 31-AUG-1999; 99DE-1041394.
XX 31-AUG-1999; 99DE-1041396.
XX 03-SEP-1999; 99DE-1042076.
XX 03-SEP-1999; 99DE-1042077.
XX 03-SEP-1999; 99DE-1042079.
XX 03-SEP-1999; 99DE-1042086.
XX 03-SEP-1999; 99DE-1042087.
XX 03-SEP-1999; 99DE-1042088.
XX 03-SEP-1999; 99DE-1042095.
XX 03-SEP-1999; 99DE-1042124.
XX 03-SEP-1999; 99DE-1042129.
XX 09-MAR-2000; 2000US-0187970.
XX
XX (BADI) BASF AG.
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-137957/14.
XX N-PSDB; AAF71886.
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in

PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 549-551; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 534 AA;
Query Match 32.5%; Score 1316; DB 22; Length 534;
Best Local Similarity 57.9%; Pred. No. 3.1e-108;
Matches 264; Conservative 59; Mismatches 123; Indels 10; Gaps 5;
Qy 15 VLQAHVDEKLDGTV-LLYIDRLHVEHTSPQAFEGRLNAGKVRPDCATLTTDHNVT 73
Db 1 VWRDHWKSGENGEPDLYIDLQLLHEVTSPQAFDGLRMTGKRLRPELHLATFDHNVT 60
Qy 74 TSKKALKDIAPIKEDD--SRTQCVTLEENKVEFGVTFGLSDKRGIVHVIPEQGFTL 131
Db 61 EGIKT-----GSLEINDKISRLQVSTLRDNCSEFGVRLHPMGDVRQGIHVTVGPQLGATQ 116
Qy 132 POTTVVCGDSHTSTHGAFAFGTSEVHVLAATQCLITKRSKNMRTQVDGELAPGV 191
Db 117 PGMTIVCGDSHTSTHGAFGMAFGTSEVHVMAQTPLPKPFTMAIEVTGELQGV 176
Qy 192 SKDVLHAIIGITAGGTGAVIEFCGSVIRSLMEARMSICNMSIEGGARAGVAPDEIT 251
Db 177 SKDLILAIITAKITGGGGQGVLEYRGEAIRKMSMDARMTWCNMSIEGARGAMIAPDQT 236
Qy 252 REYLKGRPLAPKYDSEPHKATQYKNLOSDFGAKYDIDVFIKADIVPTLTWTGSPEDV 311
Db 237 FDIYEGRENAPK--GADNDEAVAYKTLPTDEGATFDKVVIEDGSALTPTITWTGTPGQ 294
Qy 312 VPITGVVPPDPETATBAKADGRMLQYMGKAGTGMEDIPVDKVFISGTSNRIEDLRA 371
Db 295 LPLGESVSPFDTNDNDKAAAEKALQYMDLVPTGLRDKIDTVFLSGTSNRIEDLQI 354
Qy 372 AAHVKGRKAPNVKAMVPGSLVKTQAEELGDKIRPEAGFEWREAGCSMCLGNPD 431
Db 355 AADILKGHIADGMR--MMVVPSSWTWIKQEAELGLDKIFTDAGAEWRTAGCSMCLGNPD 413
Qy 432 ILAQRCASTSNRNFEGRQAGGRTMLSPVMAAA 467
Db 414 QLKPGERSAFTSNRNFEGRQPGGRTHLVSPAVAAA 449
RESULT 9
ABP39047
ID ABP39047 standard; Protein; 461 AA.
XX
AC ABP39047;
XX
XX
XX
XX 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3892.
XX
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX

PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
DR N-PSDB; ABN91592.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3892; 267pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
XX Sequence 461 AA;
Query Match 32.1%; Score 1300; DB 23; Length 461;
Best Local Similarity 54.7%; Pred. No. 6.6e-107;
Matches 258; Conservative 69; Mismatches 125; Indels 20; Gaps 5;
Qy 9 QTLVYDKVLOAHVDEKLDGTVLLYIDRLHVEHTSPQAFEGRLNAGKVRPDCATLTTD 68
Db 8 QTLFDRKWKXHLVKGEGEPQLLYIDLHLEHTSPQAFEGRLNAGKVRPDCATLTTD 67
Qy 69 HNVPTTSRKALKDIAPIKEDDSTQCVTLEENKVEFGVTFGLSDKRGIVHVIPEQG 128
Db 68 HNVPTI-----DIFN-IKDEIAHKQITLQONAKDFGVHIFDMGSDDEQGIHVHVPETG 120
Qy 129 FTLPGTIVCGDSHTSTHGAFAFGTSEVHVLAATQCLITKRSKNMRTQVDGELAP 188
Db 121 LTQPGKTVCGDSHTATGAFAGTSEVHVFAITQTLWQTKPKLNKININGSLET 180
Qy 189 GVSSKDVVLAHIGITAGGTGAVIEFCGSVIRSLMEARMSICNMSIEGGARAGVAPD 248
Db 181 GVIAXDILLYLINQYGVDFGTGYALEFTGETIKNLSMEARMTICNMAIEGAKYGLMQPD 240
Qy 249 EITFEYLKGRPLAPKYDSEPHKATQYKNLOSDFGAKYDIDVFIKADIVPTLTWTGSP 308
Db 241 ETTFDYVKGRRPYATDFDS-----SMAWKKLYSDDDAVDFDKVIELDVTNLEPQVTWTGTP 295
Qy 309 EDVVPITGVVPPDPETATBAKADGRMLQYMGKAGTGMEDIPVDKVFISGTSNRIED 368
Db 296 EMGVSFSPNPP-----EIKNANDQRAYDVMGLHPGQKAEIDIKLGYVFLGSCNTNALS 348
Qy 369 LRAAAAIVVGRKKAPNVKAMVPGSLVKTQAEELGDKIRPEAGFEWREAGCSMCLGM 428
Db 349 LIEASHIIKGQOVHNI-TAIVVPGSRTVKEAEALGLDKLPKADFEWREPGCSMCLGM 407
Qy 429 NPDILAPQRCASSTNRNFEGRQAGGRTMLSPVMAAAAGIVGLADYVKL 480
Db 408 NPDQVPEGVHCASSTNRNFEGRQAGGRTMLSPVMAAAAGIVGLADYVKL 459
RESULT 10
AAU36563
ID AAU36563 standard; Protein; 456 AA.
XX
XX AAU36563;
XX
XX 14-FEB-2002 (first entry)
DT

Db 404 NPDQVDPGVHCASTSNRNFEGROGKGARTHLVSPAMAAAAAINGHIFIDIRKV 455
RESULT 13
AAR54216
ID AAR54216 standard; Protein; 460 AA.
XX AAR54216;
XX
DT 09-NOV-1994 (first entry)
XX
DE L.lactis branched amino acid synthase leuc gene product.
XX
XX branched amino acid; ilv operon; leucine; isoleucine; valine;
KM biosynthesis; alpha-acetolactate synthase; diacetyl; food flavouring;
KM attenuation; anti-terminator; Lactococcus.
XX
OS Lactococcus lactis (subsp. lactis).
XX
XX Key Location/Qualifiers
FH Misc-difference 436
FT /note= "Val residue corresponds to CTG codon"
FT
XX
XX FR2696190-A.
XX
PD 01-APR-1994.
XX
PF 25-SEP-1992; 92FR-0011470.
XX
PR 25-SEP-1992; 92FR-0011470.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX (AGRI-) AGRIC & FOOD RES COUNCIL.
XX
PI Ehrlich S, Godon J, Renault P;
XX
XX WPI; 1994-128287/16.
DR N-PSDB; AAO64211.
XX
XX DNA coding for alpha-aceto:lactate synthase - for enhancing
PT di:acetyl prodn. in microorganisms, esp. for mfr. of dairy prods.
XX
XX
XX Disclosure; Fig 2; 45pp; French.
XX
XX The genes involved in the pathway for synthesis of branched amino
CC acids in L.lactis subsp. lactis are organised in two units
CC containing the leu (including leuc) and ilv genes, respectively.
CC Both units are necessary for the synthesis of leucine but only the
CC second unit is required for synthesis of ile and Val. The ilvB
CC and ilvN genes and the subunits of alpha-acetolactate synthase
CC that they code for are claimed.
XX
XX
SQ Sequence 460 AA;
Query Match 28.1%; Score 1138.5; DB 15; Length 460;
Best Local Similarity 50.4%; Pred. No. 1.7e-92;
Matches 240; Conservative 61; Mismatches 154; Indels 21; Gaps 6;
QY 9 QTLVYKVLQAHVYDEKLGTVLLYIDRLHVEVTSPOAFEGIRNAGRKRVRDCTLATTD 68
Db 4 KTIFFKLMDOHVIAENBEPOLLTYIDLVHIVHTSPQAFQGRAGRRVRKRDLYTGTTD 63
QY 69 HNVPTSRKALKDIASFKEKDSRTQCVTLSENVKEFGVTYFGSLDKQGIHVHVGPEQG 128
Db 64 HNVPTQNFNIODLI-----SKQIDTFTKNVKEFDVPAATHGKGGGIHMAAPBSG 116
QY 129 FTLPGTIVVCGDSHTSTHGAFGALAFGIGTSEVHVLAQTCLITKRSKNMRLQVDEGLAP 188
Db 117 RTQPEKTIIVCGDSHTATNGAFGALAFGIGTSEVHVLAQTITWQVPRKMKLIEFGHPQK 176
QY 189 GVSSEDDVULHAIIGITAGTGAVALIEFGCSVTRSLMEKRMSCIMMSIEGARGAVNAPD 248
Db 177 GIYSKDFILALIAKXVDAVGAVAYEYSGDAISDLSEMERMTICNMSTIEFGAKIGLMDP 236

QY 249 EITPEYKGRPLAEKXDSPEWHKATQYWKULQSDPGA KYDIDVEIDAKDIPVTLTWGTSP 308
Db 237 EKTVDYVKGREHAKR-----NDEAVSKEKXLVSDSDAQYDKILSLDVSQLKPMVTWGTNP 292
QY 309 EDVVPITGV-VPEDETATEKAKDGRMLQYMLKAGTPEMDIPVDKVFIFGSCNSTRIE 367
Db 293 -----GMGLEPGEKEPELINNDLNYERAYQYMDLKPQQTASDIDLGIYFISCTNARLG 345
QY 368 DLRAAAVVKGRKAPVKSAMVVPGSGLVKTOAEEGLDKIFEEAGPEMRBAGCCSMCLG 427
Db 346 DLEERAKIIGDRHIADGL-TGIIVPGSRPVKEAAEAQGLDKITFEAGFEMRPEPGCSACLG 404
QY 428 MNPDIILAPQERCASTSNRNFEGROGAGRTILMSPVMAAAAGIVGKLADVRKL-TD 482
Db 405 MNPQIPEYVHCASTSNRNFEGROGHNARTIVCSPAMAAAAAIAKGKFDVHMLVTD 460

RESULT 14
ABB54552
ID ABB54552 standard; Protein; 460 AA.
XX
XX ABB54552;
XX
XX 16-MAY-2002 (first entry)
XX
XX Lactococcus lactis protein leuc.
XX
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis IL1403.
XX
XX FR2807446-A1.
XX
XX 12-OCT-2001.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
XX WPI; 2002-043418/06.
XX
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX
XX
XX Claim 6; SEQ ID No 1254; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent parent
CC WO20017734 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 460 AA;

Query Match 27.9%; Score 1132.5; DB 23; Length 460;
Best Local Similarity 50.4%; Pred. No. 5.8e-92;
Matches 240; Conservative 59; Mismatches 156; Indels 21; Gaps 6;
QY 9 QTLVYKVLQAHVYDEKLGTVLLYIDRLHVEVTSPOAFEGIRNAGRKRVRDCTLATTD 68
Db 4 KTIFFKLMDOHVIAENBEPOLLTYIDLVHIVHTSPQAFQGRAGRRVRKRDLYTGTTD 63

```
QY 69 HNVPTTSRKALKDIAFIKEDDSRTQCVTLLENVKEFGVTYFGLSDKRQGIHVHVGPEQG 128
D 128
D 64 HNVPTQDIFNIQDLI-----SKKQIDTFTKNVKEPDVPAETHTGGKGQGIHVHVAPESG 116
QY 129 FILPGTIVVCGDSHTSTHGAFALAFAGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
D 188
D 117 RTOPGKTIIVCGDSHTATNGAFAGIAFGIGTSEVEHVLATQTIWQVKRWKIEFGQHPQK 176
QY 189 GYSSKDWLHAIGIITAGTGAVIEFCGVSIRLSMEARMSICNMSIEGGARAGWAPD 248
D 248
D 177 GYISKDFILIALIAKYGDVAGVAVESGDAISDLSEERTWCNMSIEFGAKIGLWNPD 236
QY 249 EITFEYLKGRPLAPKYDPSPEHVKATQYWKQLSDPGAKYDIDVIDAKDIPVTLTWGTS 308
D 308
D 237 EKYTYDVVKGREHAPK---NFDEAVSKWEKLVSDSDAQYKILSLDVSQKLPMTWTGTPN 292
QY 309 EDVVPITGV-VPPDETAFAEKADGRMLQYMGKLAGTGMEDIPDKVFIGSCTNSRIE 367
D 367
D 293 -----GMGLEFGKFEINNDLNYERAYQYMDLKPQOTASAIIDLGYIIFIGSCTNARLG 345
QY 368 DLRAAAVVKGRKAPNVKSAMVVPVSGLVKTOAEBEGLDKIPFEEAGFEWREAGCSMCLG 427
D 427
D 346 DLIEAAKIIGDRHIDGL-TGIVVPSRPVKEAAEAGGLDKIFKEAGFEWREPCCSACLG 404
QY 428 MNPDIILAPOERCASSTNRNPEFGQAGGRTHLMSPVNMAAAGIVGKLADVRKL-TD 482
D 482
D 405 MNPDIQIPEYVHCASSTNRNPEFGQAGGRTHLMSPVNMAAAGIVGKLADVRKL-TD 460

RESULT 15
AAW77717
ID AAW77717 standard; Protein; 264 AA.
XX AC AAW77717;
XX DT 30-OCT-1998 (first entry)
XX DE 3-Isopropylmalate dehydratase protein.
XX KW Staphylococcus aureus protein; immune response induction; eye infection;
KW antibody production; T-cell immune response; gastrointestinal infection;
KW respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW central nervous system; kidney infection; urinary tract infection;
KW antimicrobial compound identification; broad spectrum antibiotic;
KW therapy.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
FT Misc-difference 8 /note= "unspecified, encoded by TNC"
FT Misc-difference 11 /note= "unspecified, encoded by TNG"
XX FT
XX EP841394-A2.
XX PN
XX PD 13-MAY-1998.
XX PF 24-SEP-1997; 97EP-0307485.
XX XX 24-SEP-1996; 96US-0027032.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX PI Black MT, Burnham MKR, Hodgson JF, Knowles DJC, Rosenberg M;
PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;
PI Ward JM;
XX WPI; 1998-252940/23.
DR N-PSDB; AAV53507.
XX New nucleic acid sequences from Staphylococcus aureus WCHU29 -
```

```
PT useful in vaccines and for treatment of bacterial infections of e.g.
PT respiratory tract and central nervous system
XX Claim 11; Page 348-349; 390pp; English.
XX This sequence represents a Staphylococcus aureus protein, that based on
XX homology with a Lactococcus lactis subsp lactis (Streptococcus
XX lactis) protein, is a 3-isopropylmalate dehydratase (Ec 4.2.1.33)
XX (isopropylmalate isomerase) (Alpha-Ipm isomerase) (Ipmi),
XX and is encoded by a DNA sequence of the invention.
XX The DNA sequences were isolated from Staphylococcus aureus WCHU29
XX (NCIMB 40711). Host cells containing the DNA sequences are used to
XX produce polypeptides or fragments. The proteins are used in the treatment
XX of disease, for inducing an immune response by administering them, to
XX produce antibody and/or T-cell immune response. Antagonists of the
XX proteins are used for the inhibition of bacterial polypeptides.
XX Conditions which may be treated include bacterial infections, especially
XX respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
XX urinary tract, skin, bones and joints. The proteins can also be used to
XX identify antimicrobial compounds which are broad spectrum antibiotics,
XX especially useful in the treatment of H. pylori infection.
XX Sequence 264 AA;
XX Query Match 18.5%; Score 749; DB 19; Length 264;
XX Best Local Similarity 53.8%; Pred. No. 3.9e-58;
XX Matches 147; Conservative 39; Mismatches 75; Indels 12; Gaps 3;
QY 15 VLQAHVDEKLDGIVLLYIDRHLVHEVTSPOAFEGELNAGRKVRPDCATLTDHNVPTT 74
D 74
D 1 VMNRHLVXGXKQDPLLYIDLHLIHEVTSPOAFEGELNQRKLRPDLTFTATLDHNVPTI 60
QY 75 SRKALKDIASPIKEDDSRTQCVTLLENVKEFGVTYFGLSDKRQGIHVHVGPEQGTLP 134
D 134
D 61 -----DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMSDEQGIHVHVGPEGLTQPK 113
QY 135 TVVCGDSHTSTHGAFALAFAGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAPGVSSK 194
D 194
D 114 TVVCGDSHTATHGAFALAFAGIGTSEVEHVFATQTLWTKPKNLKIDINGTLPTGVYAKD 173
QY 195 VVLAHIGITAGTGAVIEFCGVSIRLSMEARMSICNMSIEGGARAGWAPDEITFEY 254
D 254
D 174 IILHLITVYGVDFGTGVALEFTGTIKNLSMDGRMTICNMAIEGGAKEYIIPDDITFEY 233
QY 255 LKGRPLAPKYDPSPEHVKATQYWKQLSDPGAKY 287
D 287
D 234 VKGRPFADNF-----AKSVDKWRRELYSDGTTRY 261

RESULT 16
AAG81974
ID AAG81974 standard; Protein; 245 AA.
XX AC AAG81974;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:1042.
XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
XX KW vaccination; endocarditis.
XX OS Staphylococcus epidermidis.
XX PN WO200134809-A2.
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US30782.
XX PR 09-NOV-1999; 99US-0164258.
XX (GLAX ) GLAXO GROUP LTD.
PA
```

XX kimmerly wj;
XX WPI: 2001-316495/33.
DR N-PSDB; AAH52824.
XX
XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
PT useful for vaccinating against infections, e.g. endocarditis -
PS Claim 18; Page 303; 2188pp; English.
XX
CC AAH5304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG61454 to AAG63120, from *Staphylococcus epidermidis*.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acid (I) may be used to produce the
CC S. epidermidis polypeptide (II) via the production of vectors
CC containing them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 245 AA;

Query Match 17.0%; Score 688.5; DB 22; Length 245;
Best Local Similarity 53.5%; Pred. No. 8.6e-53;
Matches 137; Conservative 38; Mismatches 68; Indels 13; Gaps 3;

OY 225 MEARMSICNMSIEGARAQVAPDEITFEYKGRPLAPKYDSEWHKATQYKMLQSDPG 284
DB 1 MEARMTICMAIEAGAKGLMQPDETTFNKGRYADPFS-----SMAMKELYSDD 55
OY 285 AKYDIDVFAKDIVPILTWGTSPEDEVPIITGVVDPETFAEAKKADGRMLQYMGKA 344
DB 56 AYFDKVELDVTNIEPQVTGNTPEMGVSFSPNPF-----EIKNANDORAUDYMGJHP 108
OY 345 GTPMEDIPVDVVFISGCTNSRIEDLRAAAVKGRKKAQNYKSAVVVGSGLVKTGAEE 404
DB 109 GQKADIKLGVFLGSCTNARLSDIEASHIITKGOVHPNI-TALVFGSRTVKKEAL 167
OY 405 GLDKTFEEAGFEMREAGSCMCLGMNPDLAPQERCASTSNRFEGRGAGGRTILMSFVM 464
DB 168 GLDLKLFKDGEMREPGSCMCLGMNPDPQPCVHCASSTNNFEGRGKGRATHLVSTAM 227
OY 465 AAAAGIVKLDVRL 480
DB 228 AAAAINGKFIIDVRKV 243

RESULT 17
AAB96358
ID AAB96358 standard; Protein; 423 AA.
XX
AC AAB96358;
DT 29-OCT-2001 (first entry)
XX
DE Putative 3-isopropylmalate dehydratase/acnitatease large subunit #2.
XX
KM Hyperthermophilic archaeon; hyperthermophilic protein.
XX
XX Pyrococcus abyssi.
XX

PN FR2792651-A1.
XX
XX 27-OCT-2000.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX (CNRS) CNRS CENT NAT RECH SCI.
PA (IFREMER) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Leconte O;
PI Querellou J, Weissenbach J, Saurin W, Hellig R;
XX
XX WPI: 2001-126236/14.
XX
PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode
PT proteins useful in industry -
PS Claim 7; Pages 1038-1040; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of *Pyrococcus*
CC abyssi (see AAF66431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 423 AA;

Query Match 15.7%; Score 637.5; DB 22; Length 423;
Best Local Similarity 35.3%; Pred. No. 7.3e-48;
Matches 166; Conservative 77; Mismatches 166; Indels 61; Gaps 10;

OY 10 TLVYKQVAHVDE--KLDGTVLLYIDRLVHEVTSPOAFGLRNAG-RKYVRPDCITAT 66
DB 4 TIAEKILADHSEBEVKGEIVMAKLDVFENDVTMPAIKFRBELGVKVFDRERIAIV 63
OY 67 TDHNVPTTSKRALDKDIASFIEKEDSRTOCVLEENVKEFGYTYGLSKRGQIVHVIQPE 126
DB 64 LDHFTPN-----KDI-----KSABQCKSSREFAKEMGIKWF-FEGGSVGEHCLPE 109
OY 127 QGFTLPGTTVCGPSHSTHGAFGLAFGISTSEVHEVLAITQCLITKRSKNMRLQVDEGL 186
DB 110 LGVLVPGDLITGADSHITCTYALQGFATGVSTDLAVMAATGEAMFRVPEIMKFTYBEEL 169
OY 187 AFGVSKDVLVLAIGITAGTGAIVIEFCGSVIRLSMEARMSICNMSIEGARAQVAP 246
DB 170 QPYVTGKDLIHTIGDIGNVALYKVFESGVILELSEBQRMSTNNALIEAGAKTGIIIE 229
OY 247 PDEITFEYKGRPLAPKYDSEWHKATQYKMLQSDPBAKDIVFIAKDIVPILTWGT 306
DB 230 PDKTLDYVKER-----AKRKFRVYKSDDEDAKYKYLEVVTWMEPVVAEPH 276
OY 307 SPEDVVPITGVVDPDETFATEAKKADGRMLQYMGKAKGTPMEDIPVDKVFISGCTNSRI 366
DB 277 LPENTVPI-----SKAAKK-----NKIDQVFIISGCTNSRI 307
OY 367 EDLRAAAIVKGRKKAQNYKSAVVVGSGLVKTGAEEGLDKIFEAGFEMREAGSCMCL 426
DB 308 EDLRMAAEILBGGQVAKKVR-LIVPGSPVYWKALKGLIEIFLEAGAVIGPPGICPL 366
OY 427 GMPDILAPQERCASTSNRFEGRG-AGGRTILMSPTMAAAGIVGKLA 475
DB 367 GGHMGVLASGERAVSTTRNFRVGRGHPKSEVYLANPYVAAASAVLAGRIA 416

RESULT 18
AAG29924

PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	10-SEP-1999;	99US-0152363.
PR	13-SEP-1999;	99US-0153070.
PR	15-SEP-1999;	99US-0153758.
PR	16-SEP-1999;	99US-0154018.
PR	20-SEP-1999;	99US-0154039.
PR	22-SEP-1999;	99US-0154779.
PR	23-SEP-1999;	99US-0155139.
PR	24-SEP-1999;	99US-0155486.
PR	28-SEP-1999;	99US-0155659.
PR	29-SEP-1999;	99US-0156458.
PR	04-OCT-1999;	99US-0156596.
PR	05-OCT-1999;	99US-0157117.
PR	06-OCT-1999;	99US-0157753.
PR	07-OCT-1999;	99US-0157865.
PR	08-OCT-1999;	99US-0158029.
PR	12-OCT-1999;	99US-0158232.
PR	13-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match

Best Local Similarity 13.8%; Score 558.5; DB 21; Length 461;

Matches 164; Conservative 73; Mismatches 175; Indels 99; Gaps 22.

OY	3	GAESFPQTLIDYKVLQAHVDEKL---DGTVLIIYIDRHVHEVTSPOAEFGL--RNAGRV	57
Db	13	GSVKQGMTEKIL-ARASEKSLVPGDINWNVULMTHDVCGRPAF-GIPKREFGEKA	70
OY	58	R--RDPDCLATIDHVNPTSRKALDIASFIKEDDSRQCVLNEVNEF----GVTYNG	111
Db	71	KWMDKEKIVVIDHVIIFTDKRRANRV-----DIVREHC--REONIKIFYITDLGNFK	122
OY	112	LSDKQGIHVHIVGPEOGFTLPCTTWCGDSHTSTGAFGALFAGIGTSEVEVILATOCIL	171
Db	123	ANPDYKGVCHVALAEGHCRPEVYLLGTDSHTCTAGAFQGFATIGINDAGVLTGKIL	182
OY	172	TKRSQNMRLQVNDGELAPGVSKDVVLAIGITGTAGGTGCAVIEFGGSVIRLSISMARNST	231

Db	183	LKVPFMRRIIDGEMPSVLQAKDILLQITIGELIVAGATVYKTMESGTTISLSMEERNTL	242
Qy	232	CNMSIEGARAGAVAPDEITFEYLKGRPLA---PKYDSPEWHKATQYWKNLQSDPGAKYD	288
Db	243	CNMVVEAGGKNGVIEPPDATTLNLYBACILSCFLPY-----SDGNMSEFV	286
Qy	289	IVFIFDADIVYTLTWGTSPEDEVVPIGTGVDPDETFATEAKKADGRMLQYMGAKPTM	348
Db	287	ADYRPDSVLSKEVIV---AKPHS-----PDNRLARECK-----	316
Qy	349	EDIPVDKVFISGCTNSRIEDLRAAAY--VKGK-KAPNV-----KSMVYVPGS	394
Db	317	-DVKIDRYIISGCTGCKTEDFPAAKLPHAAGRKVQFTFLVPATQKVMDVYALPVPGA	375
Qy	395	GLVKTQAEIEEGLDKIFEEAGFEW--REAGCSMCLGNPDILA---POBCASTSNRNFEG	449
Db	376	G-GKTCQAQ-----IFEEAGGCDTPASPECGACDGGPADTYARLNLEPQV-CYSTTNRNFPG	427
Qy	450	ROG-AGGRTHLMSPVMAAAGIVGLADVKK	479
Db	428	RWGHEGQIYLAAPYTAASALTGRVADPRE	458
RESULT 19			
AA	GA040226		
ID	AA040226	standard; Protein; 469 AA.	
XX	AA		
AC	AA040226;		
XX	AA		
DT	18-OCT-2000	(first entry)	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 49883.		
XX	XX		
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.		
XX	XX		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-0301439.		
XX	XX		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	23-APR-1999; 99US-0130891.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		
PR	04-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	07-MAY-1999; 99US-0132863.		
PR	11-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134421.		
PR	14-MAY-1999; 99US-0134570.		
PR	18-MAY-1999; 99US-0134768.		

us-10-010-227-3.rag

Mon Mar 17 09:02:02 2003

PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0140353.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140354.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140695.	PR	01-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	PR	07-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-0140991.	PR	10-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145313.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;		PR	26-OCT-1999;	99US-0161359.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144326.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 558.5; DB 21; Length 509;
Best Local Similarity 32.1%; Pred. No. 1.1e-40;
Matches 164; Conservative 73; Mismatches 175; Indels 99; Gaps 22;

QY 3 GAESTPOTLYDKVLQAHVDEKL---DGTVLLYIDRHLVHEVTSPOAEGL--RNAGRKV 57
Db 61 GSVKGTMTTEKIL-ARASEKSLVVPDNIWVNDVLMTHDVCVGGAF-GIFKRFEGKA 118
QY 58 R--RPDCTLATDHNVTTSRKALKDIASFIKEDDSRTQCCTVLEENVKEF-----GVTYFG 111
Db 119 KWDPEKIVIPDHYIIFTADKANRV-----DIMREHC--REQNIKIFYDITDLGNPK 170
QY 112 LSKRQGIHVHVGPEQGTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
Db 171 ANPDYKGVCHVALAQEGHCRGEVLLGTDSTCTAGAFGQFATGIGNTDAGFVLGTGKIL 230
QY 172 TKRSKNMRIQVDGELAPGVSSKDVLLHAIGIITAGGTGAVIEFCGVSIRSLSMEARMSI 231
Db 231 LKVPPTMRFIIDGEMPSYLOAKDLILQIGRISVAGATYKTMFEFSGTTIESLSMEERMTL 290
QY 232 CNMSIEGARAGMVAPEITFEVILKGRPLA---PKYDSEPHWKATQYWKNLQSDPCAKYD 288
Db 291 CNMVVEAGKNGVIPPDTATLNYVEACILSCFLPVY-----SDGNASPV 334
QY 289 IDVEIDAKDIVTLTWGTSPEDEVVPIITGVVPDPETTFATEAKKADGRMLQYMGLKAGTGM 348
Db 335 ADYRFDVSKLEPVV---AKPHS-----PDNRALARECK----- 364
QY 349 EDIPVDKVFSGCTNSRIEDLRAAAV--VKGRK-KAPNV-----KSMVVPQS 394
Db 365 -DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRVKVKPTFLVPATQKVMMDVYALPVGGA 423
QY 395 GLVKTQAEFEGLDKIFEEAGFEW-REAGSCMCLGNPDLA---POERCASSTNRNPEG 449
Db 424 G-GKTCAG-----IFEEAGCDTTPASPCGACLGCPADTYARLNEPQV-CVSTTNRNFFG 475
QY 450 RQG-AGGRTHLMSFVMAAAAGIVGLADVRK 479
Db 476 RMGHKEGQIYLASPYTAASALSGRVPADPRE 506

RESULT 21
ID ABB92924 standard; Protein; 509 AA.
AC ABB92924;
DT 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 2135.
XX
XX
XX 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 2135.
XX
XX Herbicidal; plant; agriculture; herbicide.
OS Arabidopsis thaliana.
PN W0200210210-A2.
PD 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX (FARB) BAYER AG.
PI Tietjen K, Weidler M;
XX
XX WPI; 2002-269010/31.
XX
XX
PT Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -
XX
XX
PS Claim 5; SEQ ID NO 2135; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins
(ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
useful as herbicides.
XX
SQ Sequence 509 AA;
Query Match 13.8%; Score 558.5; DB 23; Length 509;
Best Local Similarity 32.1%; Pred.No.1,1e-40;
Matches 164; Conservative 73; Mismatches 175; Indels 99; Gaps 22;
QY 3 GAESEPOLYDVLOAHVDEKLT---DGVLLYIDRHLVHEVTSQAFKGL--RNAGRKV 57
DB 61 GSVKGTMTTEKIL-ARASEKSLVPGDNIVWVDVLMTHVCGPEAF-GIFKREFGEKA 118
QY 58 R--RPDCTLATDHNVPPTSRKALDIASFIREDDSRTOCVLEENVKEF---GVTYFG 111
DB 119 KVMDEKIVIPDHIYIFNDKXANNV-----DIMREHC--REQNIKXFYVITDLGNKX 170
QY 112 LSDKRGIIVHVIPEQGFLLPQTTVCGDSRSTHGAFGALAFGIGTSEVHVLAQCI 171
DB 171 ANPDYGVCHVALAQEGHCRPGSEVLLGTDSHTCTAGAFGQFATGINTDAGVLGAKIL 230
QY 172 TKRSKMRILQVNGELAPGVSSKDVVLAIGIIGTAGGTAIVIEFQGSVRSLSMEARMSI 231
DB 231 LKVPPTMRPILDEGMPYSYQAKDLIIQIGELISVAGATYKTKMEFSGTTESSLSMERMTL 290
QY 232 CNMSEGGARAGVAPDEITFEYLKGRPLA---PKYDSEWHKATQYMKLQSDPGAKYD 288
DB 291 CNMVEAGGKGVIPDATTLVNVEACILSCLPLVY-----SDGASISV 334
QY 289 IDVFIDAKDIVLTLTWGTSPEDVVPITGVVDPDEFATKAKADGRMLQYMGAKAGTGM 348

DB 335 ADYRFDVSKLEPVV---AKPHS-----PDNRALARECK----- 364
QY 349 EDIPYKVFIGSCCTNSRIEDLRAAAAY--VKGRK-KAPNV-----KSANVFGS 394
DB 365 -DVKIDRVYIGSCGTGKTEDPMAAKLFHAAGRKVKVPTFLVPATQKVMVDVALPVGGA 423
QY 395 GLVKTQAESEGDLKIFEEAGFEW--REAGCSMCLGNPILA-----PQRCASSTNRNEG 449
DB 424 G-GKTCAG-----IFEBAGDTPASPSCAGACLGAPADTVARLNEPVV-CVSTTRNPPG 475
QY 450 RQG-AGGRTHLMSPYMAAAGIVGLADVRK 479
DB 476 RMGKKEGQIVLASPTTAASALTGRVADPRE 506
RESULT 22
ID AAG40225 standard; Protein; 517 AA.
AC AAG40225;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49882.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158212.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

13.8%; Score 558.5; DB 21; Length 517;

Match	Local Similarity	31.6%	Pred. No. 1.1e-40	Matches 164	Conservative 72	Mismatches 176	Indels 107	Gaps 22
QY	3	GAESPPQTLVYDVKVLCQAHVNDKLT---	DGTVLLYIDRHLVHEVTSPOAEGL--	RNAGRKV	57			
Db	61	GSVKGTGMNTEKIL-ARASEKSLVPGDNINWNVDLMLHDVCGGAF-GIKRFEGEKA			118			
QY	58	R-RPDCLTATTDHNVPTTSRKALKDIASFIKEDDSPTQCVTLLENKVEF---	GVTYFPG	111				
Db	119	KWMDKEKIVLVEPDHIFETADKRANRV-----	DIMREHC--REONIKKFYDITDLGNFK	170				
QY	112	LSDRKQGIHVHVGPEQGFLLPCTTVVCGDSHTSGAFALAFGIGTSEVHEVLTQGLI		171				
Db	171	ANPDYKGVCHVALAGEGHCPRGEEVLLGTDSHTCTAGAFQGFATGIGTIDAGFVLGTGKITL		230				
QY	172	TKRSKNNRIQVNDGELAPGVSSKDVVLHAIGITGATGAVIEFGSGVIRLSMEARMSI		231				
Db	231	LKVPETMRFILDGEMPSYIQAKDILLQITGELISVAGATYKIMEFGSTTESMEERTML		290				
QY	232	CNMSIEGGARAGVAPDETTFEYLK-----	GRPLAKYDSSPEWHKATQYWKQLQ	280				
Db	291	CNMVVEAGGKNGKGVIPDDATTTLLNVVACILSCFLNRTSVFPEEVY-----		334				
QY	281	SDPGAKYDIDVFIADKDIIVPLTLTWGSPEDDVPIITGVVPDEPTFAEAKKADGRMYLQYM		340				
Db	335	SDGNASFVADYRFDSVKLEPVV--AKPHS-----	PDNRLARECK-----	372				
QY	341	GLKAGTPMEDIVVDKVFISGCTNSRIEDRRAAAV--VKGRK-KAPNV-----	K	386				
Db	373	-----DVKIDKVIYIOSCTGKTEDMAAAKLFLHAGRKVKVPTFLVPATQKYMMDV		423				
QY	387	SAMVVGSGVLVYTOAEBEGLDKIFEEBAGFEW-REAGCSNCLGNMPDILA----	POERCAS	441				
Db	424	YALPVPFAG-GKTCAG-----	IFEEBAGCDTPASPSCAGCLGAPADTVARLNEPQV-CVS	475				
QY	442	TSNRNFBGRQG-AGGRTHLMSPVMAAAAGIVGKLADYRK	479					
Db	476	TTNRNFPGRMGHEGQIYLASPYTAASALGTVRADPRE	514					
RESULT 23								
XX	ID	AAAG29925	standard; Protein; 443 AA.					
XX	AC	AAAG29925;						
XX	DT	17-OCT-2000	(first entry)					
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35685.						
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;						
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;						
XX	XX	termination sequence.						
XX	OS	Arabidopsis thaliana.						
XX	PN	EP103405-A2.						
XX	PD	06-SEP-2000.						
XX	PF	25-FEB-2000; 2000EP-0301439.						
XX	PR	25-FEB-1999; 99US-0121825.						
XX	PR	05-MAR-1999; 99US-0123180.						
XX	PR	09-MAR-1999; 99US-0123548.						
XX	PR	23-MAR-1999; 99US-0125788.						
XX	PR	25-MAR-1999; 99US-0126264.						
XX	PR	29-MAR-1999; 99US-0126785.						
XX	PR	01-APR-1999; 99US-0127462.						
XX	PR	06-APR-1999; 99US-0128233.						
XX	PR	08-APR-1999; 99US-0128714.						
XX	PR	16-APR-1999; 99US-0129845.						
XX	PR	19-APR-1999; 99US-0130077.						

PR	2-MAR-1999;	99US-0130449;
PR	23-APR-1999;	99US-0130510;
PR	23-APR-1999;	99US-0130691;
PR	28-APR-1999;	99US-0131449;
PR	30-APR-1999;	99US-0132048;
PR	30-APR-1999;	99US-0132407;
PR	04-MAY-1999;	99US-0132484;
PR	05-MAY-1999;	99US-0132485;
PR	06-MAY-1999;	99US-0132486;
PR	06-MAY-1999;	99US-0132487;
PR	07-MAY-1999;	99US-0132663;
PR	11-MAY-1999;	99US-0134256;
PR	14-MAY-1999;	99US-0134219;
PR	14-MAY-1999;	99US-0134221;
PR	14-MAY-1999;	99US-0134370;
PR	18-MAY-1999;	99US-0134768;
PR	19-MAY-1999;	99US-0134941;
PR	20-MAY-1999;	99US-0135424;
PR	21-MAY-1999;	99US-0135553;
PR	24-MAY-1999;	99US-0136021;
PR	25-MAY-1999;	99US-0136629;
PR	27-MAY-1999;	99US-0136592;
PR	28-MAY-1999;	99US-0136782;
PR	01-JUN-1999;	99US-0137222;
PR	03-JUN-1999;	99US-0137528;
PR	04-JUN-1999;	99US-0137502;
PR	07-JUN-1999;	99US-0137794;
PR	08-JUN-1999;	99US-0138024;
PR	10-JUN-1999;	99US-0138540;
PR	10-JUN-1999;	99US-0138847;
PR	14-JUN-1999;	99US-0139119;
PR	16-JUN-1999;	99US-0139452;
PR	16-JUN-1999;	99US-0139453;
PR	17-JUN-1999;	99US-0139492;
PR	18-JUN-1999;	99US-0139454;
PR	18-JUN-1999;	99US-0139455;
PR	18-JUN-1999;	99US-0139456;
PR	18-JUN-1999;	99US-0139457;
PR	18-JUN-1999;	99US-0139458;
PR	18-JUN-1999;	99US-0139459;
PR	18-JUN-1999;	99US-0139460;
PR	18-JUN-1999;	99US-0139461;
PR	18-JUN-1999;	99US-0139462;
PR	18-JUN-1999;	99US-0139463;
PR	24-JUN-1999;	99US-0140623;
PR	28-JUN-1999;	99US-0140625;
PR	29-JUN-1999;	99US-0140591;
PR	30-JUN-1999;	99US-0141847;
PR	01-JUL-1999;	99US-0141848;
PR	01-JUL-1999;	99US-0142154;
PR	01-JUL-1999;	99US-0142055;
PR	02-JUL-1999;	99US-0142505;
PR	06-JUL-1999;	99US-0142390;
PR	15-JUL-1999;	99US-0144005;
PR	16-JUL-1999;	99US-0144085;
PR	16-JUL-1999;	99US-0144086;
PR	19-JUL-1999;	99US-0144325;
PR	19-JUL-1999;	99US-0144331;
PR	19-JUL-1999;	99US-0144332;
PR	19-JUL-1999;	99US-0144333;
PR	19-JUL-1999;	99US-0144334;
PR	19-JUL-1999;	99US-0144335;

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 550.5; DB 21; Length 443;
Best Local Similarity 32.1%; Pred. No. 4.6e-40;
Matches 162; Conservative 72; Mismatches 171; Indels 99; Gaps 22;

QY 10 TLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPOAEGE--RNAGRKVR--RPDC 62
Db 2 TWTEKIL-ARASEKSLVVPFGDNIWVVDVLMTHDVCGGAP-GIFKRFEGEKAKVWDPEK 59

QY 63 TLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENKVF-----GVTYFGLSKROG 118
Db 60 IIVIPDHVIFTADKANRV-----DIMREHC--REQNIKYFYDITDLGNFKANPDYKG 111

QY 119 IVHVIGPEQGFILPGTTVWCDSHTSHGAFGALAFGIGTSEVHVLAQCLITKSKNM 178
Db 112 VCHVALAQEGHCRPEVLLGTDSTCTAGAFGQFATGNTDAGFVLGTGKILLKVPPTM 171

QY 179 RIOVDGELAPGVSKDVLVHAIGIAGTGAGTGAVIEFCGVSIRLSMEARMSICNNISIEG 238
Db 172 RFILDGEMPSYLOAKDLILOIIGISVAGATYKTFEFGTIESLSMEERTLCNNVVEA 231

QY 239 GARAGMVAPDEITFEYLKGRPLA---PKYDSPENHKATQYWNKLOSDPGAKYDIDVFIDA 295
Db 232 GKGNGVIPPDPATTLNYYEACILSCFLPVY-----SDGNASFVADYRFDV 275

QY 296 KDIVPTLTWGTSPEDVVPITGVVDPDPETFAEAKADGRMLQYMGKAGTIPMEDIPVDK 355
Db 276 SKLEPVV---AKPHS-----PDNRALARECK-----DVKIDR 304

QY 356 VFICSGCTNSRIEDLRAAAV--VKGRK-KAPNV-----KSAMVVPGSLVKTOA 401
Db 305 VYIGSCGTGGKTEDPMAAKLFLHAAGRKVKVFTFLVPATQKVMWDVYALVPAG-KTCA 363

QY 402 EEEGLDKIFEAGFEW-REAGCSMCLGMNPDILA-----POERCASSTNRNFEGRQG-AGG 455
Db 364 Q-----IFEEAGCDTPASPCGACLGPGADTYARLNPEQV-CVSTTNRNFFGRMGHKEG 416

QY 456 RTHLMSPVMAAAGIVGKLADVRK 479
Db 417 QIYLASPYTAAASALTGRVADPRE 440

RESULT 24
AAG40227
ID AAG40227 standard; Protein; 451 AA.
XX
AC AAG40227;
XX
DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 49884.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126254.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132663.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

```

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 550.5; DB 21; Length 451;
Best Local Similarity 31.6%; Pred. No. 4.7e-40;
Matches 162; Conservative 71; Mismatches 172; Indels 107; Gaps 22;

QY 10 TLYDKVLQAVHDEKL---DGTULLYIDRLHVEHTSPQAFGL--RNAGEKVR--RPDC 62
DB 2 TWTEKIL-ARASEKSLVPGDNVWVDVLMTHDVCQPGAF-GIFKREFGSEKAKWDPK 59
QY 63 TLATTNNVPTTSRKALKDIASFIKEDSRTQCVTLLENVKEF---GVTFYGLSDKRQ 118
DB 60 IVVIPHDIYFADKRRANRV-----DIMREHC--REQNIKYFDITDLGNFKANPDYKG 111
QY 119 IVHVIGPEQOFTLPGTTVVCGDSTHTGAFGALAFGIGTSEVBEHVLAQCLITKRSKNM 178
DB 112 VCHVALAQEGHCRPGEVLLGTDSTCTAGAFGQFATGNTDAGFVLGTGKILLKVPPTM 171
QY 179 RIOVDGELAPGVSKQVWLHAIGLITAGGTGAVIEGSGVIRSLSEARMSICNMSTEG 238
DB 172 RFILDGEMPYSLOAKDILQIGISVAGATYKTMESFGTIESLSMEERTLCNMVVEA 231
QY 239 GARAGWVAPDEITTEYLK-----GRPLAPKYDSEPHWKATQYKKNLQSDPDGAKY 287
DB 232 GGNKVGIPTTATTLNLYVEACILSFLNRTSVFPFVY-----SDGNASF 275

```

```

QY 288 DIDVFDIAKDIVPTLTWTGTSPEVDVITGVVPDETPATEAKKADGRRLQYMGKAGCTP 347
DB 276 VADYREFDVSKLEPVV---AKPHS-----PDNRALARECK----- 306
QY 348 MEDIPVDKVFISGCTNSRTIEDLRAAAV--VKGRK-KAPNV-----KSAMVVPG 393
DB 307 --DVKIDRVYIGSTGKTEDFMAAKLPHAAGRKVKVPTFLVPATQKVMVDVVALPVG 364
QY 394 SGLVKTQABEEGLDKIFEAGFEW-REAGSCMLGMNPDILA-----POERCASSTNRNFE 448
DB 365 AG-GKCAQ-----IFEEAGCDTPASPCGACILGGPADTYARLNEPQV-CVSTTRNRP 416
QY 449 GROG-AGGRTHLMSPVMAAAGIVGKLADVRK 479
DB 417 GRMGHKEGQIYLASPYTAAASALTGRVADPRE 448

RESULT 25
AAU36307
ID AAU36307 standard; Protein; 212 AA.
XX
AC AAU36307;
XX
XX 14-FEB-2002 (first entry)
DT
DE Pseudomonas aeruginosa cellular proliferation protein #297.
XX
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
XX WO200170955-A2.
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR
DR N-PSDB; AAS54166.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11900; 51lpp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an

```

CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 212 AA;

Query Match 13.4%; Score 542; DB 22; Length 212;
Best Local Similarity 53.3%; Pred. No. 8e-40;

Matches 112; Conservative 34; Mismatches 50; Indels 14; Gaps 6;

QY 543 FTILKGIAPPEKANVDDAIIPKQFLKTKTGIGNALFYEMRF--NEDGTEKS--- 595
DB 4 YTOTGLVAPLDIRANVDDQIIPKQFLKSIKRTGFGPNLFDWMRYLDVGQPGQDNKRPL 63
QY 596 --DEVLNKEPYRKASILVCTGANFGCGSSREHAPWALNDFGIRSVTAPSFADIPFNNSPK 653
DB 64 NPDPVLANQPRYQASVLIAR-ENFGCGSSREHAPWALNDFGIRSVTAPSFADIPFNNSPK 122
QY 654 NGMLPIPIKDOAQIEAI--AAEARAGKEIEVDLPNQLIKNATGETICTFEVEEPRKHCLV 711
DB 123 NGLLPI-ILPEAEVDELFRQVEANEGYQLSIDLAQVTRPDGK-VLGFEVDPRKHCLL 180
QY 712 NGDDDIGLTMQMEDKIAFEFAKMTRETPWL 741
DB 181 NGDDDIGLTLQDADAIRAFEDGYRQQQPWL 210

Search completed: March 17, 2003, 08:49:07
Job time : 53 secs

